NAMI - 1047

DEVELOPMENT OF AN AVIATION COMBAT CRITERION: PRELIMINARY REPORT

George M. Rickus, Jr. and James R. Berkshire

August 1968

NAVAL AEROSPACE MEDICAL INSTITUTE
NAVAL AEROSPACE MEDICAL CENTER
PENSACOLA, FLORIDA

DEVELOPMENT OF AN AVIATION COMBAT CRITERION: PRELIMINARY REPORT

George M. Rickus, Jr. and James R. Berkshire

Bureau of Medicine and Surgery
M4305.09-3002.1

Approved by

Ashton Graybiel, M. D.
Head, Research Department

Released by

Captain J. W. Weaver, MC, USN
Commanding Officer

6 August 1968

NAVAL AEROSPACE MEDICAL INSTITUTE
NAVAL AEROSPACE MEDICAL CENTER
PENSACOLA, FLORIDA 32512

SUMMARY PAGE

THE PROBLEM

An adequate criterion of combat performance is a necessity if the Naval Aviation Training Command is to have proper feedback concerning its aviator and flight officer output. This paper considers the various approaches to the development of a combat criterion and describes the methodology and preliminary findings of the one approach judged to be optimal.

FINDINGS

The preliminary results indicate that the flight surgeon is capable of identifying unsatisfactory combat pilots and flight officers by using the following criteria: "turned in wings," "had wings taken away," "transferred due to poor performance," "given non-flying duties," or "nominated as person others refuse to fly with." The training performance of these unsatisfactory men was below average. It appears that only Peer Rating has possible value as a screening device to prevent potentially inadequate performers from reaching the fleet.

INTRODUCTION

The problem of developing an adequate criterion of combat performance within aviation has plagued aviation psychologists since the beginning of World War II. A survey of the literature reveals both practical and technical difficulties. Jenkins and associates during World War II reported that when data were collected in a combat theater, there were transportation difficulties, lack of cooperation by some commanding officers, inconsistency in methodology due to rapidly changing circumstances beyond the control of the investigators, and a lack of acceptance of the investigators (1).

During the Korean War, some criterion studies focused on specific aspects of missions, but the data reflected the inadequacies of objective records of combat performance. Ellis found essentially zero between-mission reliability for scores derived from gun-camera records (2). Hemphill and Sechrest reported that the reliability of judgments on combat bombing accuracy as recorded by strike photographs was not significantly different from zero (3).

In general, previous studies of a combat criterion have indicated some success with subjective judgment techniques, as exemplified by the work of Jenkins (1) and of Flanagan (4). This has led to a reaction in combat performance research against atomistic assessment in favor of holistic evaluation (5). The reason for this can best be understood when one considers the multidimensionality or complexity of combat flying. If one measures performance on only one dimension of combat flying, then he can only make evaluations of that particular dimension. An adequate combat criterion must necessarily reflect the many dimensions of combat flying and appropriately weight them to arrive at a final index of proficiency. To date only subjective evaluations have been able to do this with any degree of success.

The purpose of this study was to develop an adequate criterion of combat performance in order that the Naval Air Training Command might have proper feedback concerning its pilot and flight officer* (NFO) product. This study discusses the preliminary findings.

PROCEDURE

Interviews were conducted with 40 pilots and NFO's who had recently returned from Viet Nam. The purpose of these interviews was to discuss the usefulness of subjective evaluations in identifying unsatisfactory pilots or NFO's. These Viet Nam veterans indicated that peer evaluations take place informally among squadron members in most squadrons throughout a deployment. These evaluations often result in common agreement as to who is unsatisfactory, and, in most cases, these men are characterized as, "men

---

* Flight Officer (NFO) - performs the duties of navigator, radar intercept officer, et cetera.

others refuse to fly with." Other expressions used to describe these aviators were: "turned in wings," "had wings taken away," "transferred from squadron due to poor performance," and "removed from flying duties due to poor performance."

In discussing who would be in the best position to identify such men for research purposes there was general agreement on the flight surgeon as the man. The interviewed pilots and NFO's indicated that the flight surgeon knows about the informal peer evaluations because he is an accepted member of the squadron and is close to the squadron members throughout the deployment. He also keeps records on flight disposition boards and any other change of flight status.

Flight surgeons who had recently returned from combat squadrons were then interviewed. In all cases these flight surgeons readily identified to the interviewer unsatisfactory pilots and NFO's who were characterized by one or a combination of the descriptive phrases previously mentioned. From the results of these interviews, it was decided to send questionnaires to all combat deployed flight surgeons, requesting that they identify unsatisfactory pilots or NFO's, using the previously mentioned characteristics as their criteria. They were required to state the reason why the man was so identified in each case.

After an unsatisfactory group was identified, a comparison group was constructed by including the next man in the student records who went through the same advanced training syllabus, at the same time, as the man identified as unsatisfactory. Selecting by syllabus type in this way controlled for known differences in training performance between pilots of helicopters, jets, and propeller aircraft and naval flight officers. Tests for the significance of mean differences were used to compare the two groups on 17 selection and training variables.

Mean differences between groups, although informative, indicate little about the possible usefulness of a measure as a screening device. Our objective is to eliminate men from aviation training who are likely to perform unsatisfactorily in the fleet; therefore, we must focus on identifying these potential fleet unsatisfactory men by low grades on selection or training measures. To this end, frequency distributions on the same 17 variables were also analyzed to determine if and where minimum standards might best be applied.

RESULTS

One hundred and seven questionnaires were mailed to combat deployed flight surgeons; however, 33 were not received due to address problems. Of the 74 received by the flight surgeons, 44 were returned in useable form. This represents a 57% return rate. From these questionnaires, 100 unsatisfactory men were identified. Twenty-eight of these men were not included in this analysis, however, due to nonavailability of their training records.

Table I contains the results of a content analysis of the reasons given by the flight surgeons for identifying these men as unsatisfactory in combat.

Table I

Content Analysis of the Reasons Given for Identifying Unsatisfactory Aviators

N	Reason
38	Nominated as a person others refuse to fly with
36	Turned in his wings
12	Had his wings taken away from him
7	Transferred out of the squadron due to poor performance
7	Given nonflying duties within the squadron due to poor performance

Tests for the significance of mean difference showed that the unsatisfactory group had significantly lower mean performance scores than the comparison group on four of the 17 training variables studied: Pre-Flight final grade--a composite of academic course grades taken prior to actual flight training; Basic Flight grade--a composite of the phase grades within this stage; Advanced Ground School grade--another academic composite grade; and Advanced Flight Grade--a composite of the phase grades within this stage. It is of interest to note that, on 15 of the 17 variables, the unsatisfactory group had lower mean scores than the comparison group (Table II).

Frequency distributions of the 17 selection and training variables were examined for possible minimum cut-off scores that would reflect a practical training elimination ratio between those unsatisfactory pilots and NFO's in combat and those who were satisfactory. Before this could be accomplished, however, the comparison group was adjusted so that it would be more representative of the actual population. An assumption was made that 90% of the combat pilots and NFO's are satisfactory; therefore, each one in this group was multiplied by nine. After this adjustment was made, only one measure appeared promising for the application of a minimum standard.

On Peer Rating, obtained in the eighth week of pre-flight training, a minimum standard of 35 would have eliminated four men from the total sample. All four of these men were identified as unsatisfactory (Table III). This finding supports the minimum standard of 35 advocated by Berkshire in his fleet follow-up study of fiscal 1956 Naval Air Training graduates (6).

Table II

Mean Performance Scores on 17 Selection and Training Variables for the Unsatisfactory and Comparison Aviction Groups

Selection and Training Measures	Comparison Group			Unsatisfactory Group		
	M	S. D.	N	M	S. D.	N
1. Aviation Qualification Test	82.73	13.63	70	83.37	11.70	63
2. Mechanical Comprehension Test	60.38	8.48	70	62.73	7.44	63
3. Spatial Apperception Test	21.23	5.56	70	22.30	4.95	63
4. Biographical Inventory	39.65	12.75	70	38.48	12.22	63
5. Flight Aptitude Rating	6.24	1.50	70	6.54	1.33	63
6. Peer Rating	49.73	10.13	40	52.38	10.48	42
* 7. Pre-Flight Final Grade	50.81	4.75	72	52.51	5.22	63
8. Pre-Solo	3.03	0.10	58	3.03	0.05	59
9. Precision	3.03	0.10	64	3.05	0.11	62
10. Field Carrier Qualification	2.95	0.05	35	2.96	0.06	37
11. Carrier Qualification	2.95	0.15	54	2.96	0.17	60
*12. Basic Flight	2.99	0.06	57	3.01	0.06	61
13. Basic Ground	49.07	5.70	59	49.76	7.29	64
14. Basic Officer-Like-Qualities	33.77	4.90	25	34.26	6.22	26
*15. Advanced Flight	2.98	0.06	56	3.00	0.06	58
*16. Advanced Ground	49.35	4.64	56	51.52	4.85	58
17. Advanced Officer-Like-Qualities	32.46	1.78	21	32.98	1.87	23

* Difference significant at the .05 level for a two-tailed test.

Table III

Frequency Distributions of Peer Rating Scores

Score Intervals	Unsatisfactory Group	Satisfactory Group
35 or less	4	0
36 - 40	2	45
41 - 45	6	54
46 - 50	6	72
51 - 55	8	63
56 - 60	4	90
61 - 65	3	27
66 or more	3	18

CONCLUSIONS

The preliminary results indicate that the flight surgeon is identifying as unsatisfactory a group whose training performance was below average. The fact that there are aptitude and training performance differences between men identified as unsatisfactory and those not so identified increases confidence in the validity of the flight surgeon's report. It appears from the preliminary data that only Peer Rating has possible value as a screening device to prevent potentially inadequate performers from reaching the fleet.

The major problem experienced in the collection of the preliminary data was the rate of questionnaire return. For a criterion to be most effective the data must be collected on a routine basis, include all combat aviators, and contain all relevant information on performance. It appears that so far we have been only partially successful. Therefore, in order to improve the criterion data, visits will be made to combat squadrons to discuss with the flight surgeons a revised system for the routine transmittal of criterion data.

REFERENCES

1. Jenkins, J. G., The combat criterion in naval aviation. NRC Committee on Aviation Psychology, Report No. 6. Washington, D. C.: Division of Aviation Medicine, Bureau of Medicine and Surgery, U. S. Navy, 1956.

2. Ellis, D. S., Gun camera records as measure of pedestal sight gunnery proficiency. AFPRTC TN 56-30. Lackland Air Force Base, Texas: USAF Personnel Training Research Center, 1956.

3. Hemphill, J. K., and Sechrest, L. B., A comparison of three criteria of aircrew effectiveness in combat over Korea. J. Appl. Psychol., 36: 323-327, 1952.

4. Flanagan, J. (Ed.), The aviation psychology program in the Army Air Forces. Report No. 1. Pittsburgh, Pennsylvania: University of Pittsburgh, 1948.

5. Krendel, E. S., and Bloorm, J. W., The natural pilot model for flight proficiency evaluation. NAVTRADEVCEN TR 323-1. Port Washington, New York: U. S. Naval Training Device Center, 1963.

6. Berkshire, J. R., Minimum standards for naval air training: Final report on the fleet follow-up study of fiscal 1956 graduates. Special Report 58-31. Pensacola, Fla.: Naval School of Aviation Medicine, 1958.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Naval Aerospace Medical Institute
Pensacola, Florida 32512

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
DEVELOPMENT OF AN AVIATION COMBAT CRITERION: PRELIMINARY REPORT.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)
George M. Rickus, Jr., LT, MSC, USNR
James R. Berkshire

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6 August 1968	7	6

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
M4305.09-3002	NAMI-1047
b. PROJECT NO	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	1

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This is a preliminary report on the development of a combat criterion for naval aviators and flight officers. Interviews with veteran combat aviators, flight officers, and flight surgeons indicated that the flight surgeon was the single best individual capable of identifying unsatisfactory combat personnel. The criteria for identifying a man as unsatisfactory were: "turned in wings," "had wings taken away," "transferred due to poor performance," "given nonflying duties," or "nominated as person others refuse to fly with." Questionnaires were sent to all combat deployed flight surgeons, and 57% were completed and returned at the time of this analysis. Results indicate that the flight surgeon identified as combat unsatisfactory that group of officers whose aviation training performance had been below average. Frequency distributions of 17 selection and training variables indicated that only Peer Rating has possible value as a screening device to prevent potentially inadequate performers from reaching the fleet.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6801

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Criterion development						
Combat performance						
Aviation training						
Performance evaluation						
Selection						

AD675214

ERRATA TO

Rickus, Jr., G. M., and Berkshire, J. R. Development of an aviation combat criterion preliminary report. NAMI - 1047. Pensacola, Fla.: Naval Aerospace Medical Institute, 6 August 1968.

CORRECTION: Replacement of Table II, page 4.

Table II

Mean Performance Scores on 17 Selection and Training Variables for the Unsatisfactory and Comparison Aviation Groups

Selection and Training Measures	Unsatisfactory Group			Comparison Group		
	M	S. D.	N	M	S. D.	N
1. Aviation Qualification Test	82.73	13.63	70	83.37	11.70	63
2. Mechanical Comprehension Test	60.38	8.48	70	62.73	7.44	63
3. Spatial Apperception Test	21.23	5.56	70	22.30	4.95	63
4. Biographical Inventory	39.65	12.75	70	38.48	12.22	63
5. Flight Aptitude Rating	6.24	1.50	70	6.54	1.33	63
6. Peer Rating	49.73	10.13	40	52.38	10.48	42
* 7. Pre-Flight Final Grade	50.81	4.75	72	52.51	5.22	63
8. Pre-Solo	3.03	0.10	58	3.03	0.05	59
9. Precision	3.03	0.10	64	3.05	0.11	62
10. Field Carrier Qualification	2.95	0.05	35	2.96	0.06	37
11. Carrier Qualification	2.95	0.15	54	2.96	0.17	60
*12. Basic Flight	2.99	0.06	57	3.01	0.06	61
13. Basic Ground	49.07	5.70	59	49.76	7.29	64
14. Basic Officer-Like-Qualities	33.77	4.90	25	34.26	6.22	26
*15. Advanced Flight	2.98	0.06	56	3.00	0.06	58
*16. Advanced Ground	49.35	4.64	56	51.52	4.85	58
17. Advanced Officer-Like-Qualities	32.46	1.78	21	32.98	1.87	23

* Difference significant at the .05 level for a two-tailed test.